California
Commission on Teacher Credentialing

Meeting of
December 3-4, 2003

AGENDA ITEM NUMBER: **PERF - 1**

COMMITTEE: **Performance Standards Committee**

TITLE: **Recommended Passing Standards for the California Subject Examinations for Teachers (CSET): Science (Specialized) Subtests IV in Biology, Chemistry, Earth and Planetary Science, and Physics**

__X__ Action

_____ Information

_____ Report

Strategic Plan Goal(s):

Goal 1: **Promote educational excellence through the preparation and certification of professional educators**

- Sustain high quality standards for the preparation of professional educators
- Sustain high quality standards for the performance of credential candidates

Presented By: **Diane Tanaka, Phil Fitch**

Prepared By: _____ **Date:** _____
Diane Tanaka
Assistant Consultant, Professional Services Division

Prepared By: _____ **Date:** _____
Phil Fitch, Ph.D
Consultant, Professional Services Division

Approved By: _____ **Date:** _____
Amy Jackson
Administrator, Professional Services Division

Approved By: _____ **Date:** _____
Beth Graybill
Interim Director, Professional Services Division

Authorized By: _____ **Date:** _____
Dr. Sam W. Swofford
Executive Director

Recommended Passing Standards for California Subject Examinations for Teachers (CSET) Science Subtests IV in Biology, Chemistry, Earth and Planetary Science, and Physics

Professional Services Division

December 3-4, 2003

Executive Summary

The purpose of this report is to provide the Commission with recommendations, based on the informed judgments of California educators, relevant to the determination of passing standards for the CSET: Science (Specialized) Subtests IV in biology, chemistry, Earth and planetary science, and physics.

In April 2003, the Commission adopted a policy for establishing new single subject credentials for Science (Specialized) in the four science areas of biology, chemistry, Earth and planetary science, and physics. Since then, Commission staff has worked with the CSET Science Advisory Panels and National Evaluation Systems, Inc. (NES®) to develop the subject matter examinations for these new credentials. It was determined that for each of the new credentials, the examination would consist of the current CSET: Science Subtest III in the specific science area, and a new CSET: Science (Specialized) Subtest IV which would cover the necessary general science content knowledge specific to each of the science areas.

The CSET Science (Specialized) Subtests IV in biology, chemistry, Earth and planetary science, and physics were first administered on September 20, 2003. On October 15-16, 2003 and November 12-13, 2003, Commission staff and NES conducted standard setting studies for the new examinations. This report describes the standard setting studies and the results of the studies, and provides staff-recommended passing standards.

With the adoption of passing standards for these subtests, the Commission will have established an examination route for verification of the subject matter requirement for the Single Subject Science (Specialized) credentials in the four science areas of biology, chemistry, Earth and planetary science, and physics.

Fiscal Impact Summary

What passing standards should be established for the CSET: Science (Specialized) Subtests IV in biology, chemistry, Earth and planetary science, and physics?

Policy Issues to be Considered

NES is developing the CSET at no cost to the Commission; the contractor will be compensated directly from examinee fees. This test development work includes the standard setting studies described in this report.

Recommendation

Staff recommends that the Commission adopt the proposed passing standards found on page 20 of this report for of the CSET: Science (Specialized) Subtests IV in biology, chemistry, Earth and planetary science, and physics.

Recommended Passing Standards for California Subject Examinations for Teachers (CSET): Science (Specialized) Subtests IV in Biology, Chemistry, Earth and Planetary Science, and Physics

Professional Services Division

December 3-4, 2003

Overview of the Report

This report describes the standard setting studies for the CSET: Science (Specialized) Subtests IV in biology, chemistry, Earth and planetary science, and physics, and provides staff-recommended passing standards for each subtest for use in the establishment of the subject matter examination for the Single Subject Credentials in Science (Specialized). Part I provides background information on the development of the CSET: Science (Specialized). Part II describes the standard setting procedures used and the results of the studies. Part III includes staff recommendations for action.

Part I: Background Information

The Subject Matter Competence Requirement for Single Subject Teaching Credentials in Science (Specialized)

Teacher candidates in California are required to demonstrate competence in the subject matter they will be authorized to teach. Single Subject teacher candidates have two options available for satisfying this requirement. Currently, they can either complete a Commission-approved subject matter preparation program or they can pass the appropriate Commission-adopted subject matter examination(s). Because they satisfy the same requirement, these two options are as aligned and congruent as possible.

In April 2003, four new Single Subject Teaching Credentials in Science (Specialized) were established in the science areas of biology, chemistry, Earth and planetary science, and physics. The Single Subject Teaching Credential in Science (Specialized) authorizes instruction only in the specific science area listed on the credential in K-12 California public schools, and does not authorize teaching general, introductory, or integrated science. For example, a teacher who holds a Single Subject Teaching Credential in Chemistry (Specialized) would be authorized to teach chemistry classes only - if s/he were assigned to teach chemistry as well as other general or integrated science classes, a full Single Subject Teaching Credential in Chemistry would be required.

These new credentials allow candidates with advanced degrees in a specialized science field, who decide, as career changers, to enter the teaching profession. Their subject matter preparation has already been demonstrated in their chosen field of study through their advanced degrees, albeit in a single science discipline. Allowing individuals to earn a specialized science authorization in

their field of expertise provides additional flexibility for those considering a career as a science teacher and provides flexible staffing options for districts and schools who currently have difficulty finding credentialed teachers in science.

Demonstrating subject matter competency under NCLB requires a subject matter examination, or a major/degree or its equivalent in the specific subject area of instruction. The Commission previously approved three options for prospective teachers to fulfill the subject matter competency requirement for the new Single Subject Teaching Credentials in Science (Specialized), which are consistent with the NCLB teacher requirements. The options are:

- Passage CSET Subtests III and IV for the science specialty area (i.e. biology, chemistry, Earth and planetary science, or physics);
- Completion of a post-baccalaureate degree program from a regionally-accredited institution in the specialized science area, or in a closely related area deemed equivalent by the Commission; or
- Completion of a baccalaureate degree, and 30 semester units or 45 quarter units of postgraduate coursework in the in the specialized science area, or in a closely related area deemed equivalent by the Commission. The degree and postgraduate coursework must be from a regionally-accredited institution, and each course used to meet the subject matter requirement of the specialized credential must have grades of B or better, "pass" or "credit".

Development of Subject Matter Requirements for
Single Subject Teaching Credentials in Science and Science (Specialized)

In January 2001, the Executive Director appointed the Science Subject Matter Advisory Panels in the areas of biology, chemistry, Earth and planetary science, and physics to advise Commission staff on the development of the new science subject matter program standards and examinations for the Single Subject Teaching Credentials in Science. Each panel consisted of:

- classroom teachers of the subject area,
- subject area specialists in school districts, county offices of education, and postsecondary institutions,
- professors in the subject area teaching in subject matter preparation programs,
- teacher educators,
- members of relevant professional organizations,
- members of other relevant committees and advisory panels, and
- a liaison from the California Department of Education.

In April 2001 the Executive Director signed a contract with the American Institutes for Research (AIR) to work with Commission staff and the advisory panels to develop and validate, for English, mathematics, sciences and social science, subject matter requirements (SMRs) for prospective secondary teachers. With leadership from Commission staff and assistance from AIR staff, the Science Subject Matter Advisory Panels developed the science SMRs--the subject-specific knowledge, skills, and abilities needed by beginning teachers of biology, chemistry, Earth and planetary science, and physics. In May of 2001, the Commission adopted Subject Matter Requirements for Science (for more information see

http://www.ctc.ca.gov/aboutctc/agendas/june_2002/June_2002_PERF-1.pdf).

The science SMRs included the following content areas:

- General Science
- Biology
- Chemistry
- Earth and Planetary Science
- Physics

The science SMRs included one shared set of general science knowledge, skills, and abilities in which the four science concentrations are represented and, a separate set of knowledge, skills, and abilities specific to each specialized science field (biology, chemistry, Earth and planetary science, and physics).

In March of 2002, National Evaluation Systems, Inc. (NES®) was awarded the contract for the continued development and administration of the new CSET program. After the Commission's April 2003 adoption of the recommendation to establish four new Single Subject Teaching Credentials in Science (Specialized), Commission staff and the Science Subject Matter Advisory Panels began working with NES staff in the development of SMRs for the new science credentials. Since the new credentials only authorized instruction in the specific science area listed (i.e., biology, chemistry, Earth and planetary science, and physics), the science (specialized) SMRs for each of these content areas would be the associated SMRs previously approved by the Science Subject Matter Advisory Panels.

The CSET Examinations for the Science (Specialized) Credentials

Once the subject matter requirements specific to the identified science concentration for the Single Subject Teaching Credentials in Science (Specialized) were established, test development began in the Spring of 2003. Because the new credential area was "specialized", i.e., relegated to a particular science area, there was no requirement that the candidate demonstrate content knowledge in the other areas of science (as included within General Science Subtests I and II). The Science Subject Matter Advisory Panels working in conjunction with Commission staff determined that the new examination requirement would consist of two subtests for each of the science areas. The recommended structure of the CSET examination for the Science (Specialized) Credential in biology, chemistry, Earth and planetary science, and physics is as follows:

- the existing CSET: Science Subtests III the test that measures content knowledge in the specialized science field, and
- new Subtests IV that are developed from the appropriate parts of the General Science Subtests I and II of the CSET: Science examinations. The new CSET: Science (Specialized) Subtests IV consist of the general science knowledge, i.e., the foundational content knowledge, specific to each of four science areas.

All test items for the new Subtest IV (with the exception of chemistry constructed-response items) were drawn from the existing CSET: Science Subtest I and II item banks. These item banks consist of all Subject Matter Advisory Panel-approved multiple-choice and constructed-

response items. Before being added to the CSET test item bank, all newly developed items have been: 1) reviewed by the CSET Bias Review Committee; 2) reviewed and revised (if needed) by Subject Matter Advisory Panel members who are credentialed or degree holders within the science field; and 3) field tested (after the advisory panel and Bias Committee reviews). Additionally, marker responses for the chemistry constructed-response items newly developed for the CSET: Science (Specialized) Subtest IV have been selected with the assistance of the Chemistry Subject Matter Advisory Panel members.

These paper-and-pencil tests consist of both multiple-choice and "focused" constructed-response (or essay) items. The "focused" constructed-response items are scored using a three-point scoring scale. Refer to the Appendix A for the constructed-response performance characteristics and scoring scales. For comparison purposes, the subtest structures of both the original CSET: Science and CSET: Science (Specialized) examinations are listed in Tables 1 and 2 on the next page.

Table 1: Subtest Structure of the Original CSET: Science

CSET: Science Subtests	Number of Multiple-Choice Items per Test Form	Number of Constructed-Response Items per Test Form
I: General Science: Astronomy; Dynamic Processes of the Earth; Earth Resources; Waves; Forces and Motion; Electricity and Magnetism	58	2 (focused)
II: General Science: Ecology; Genetics and Evolution; Molecular Biology and Biochemistry; Cell and Organismal Biology; Heat Transfer and Thermodynamics; Structure and Properties of Matter	58	2 (focused)
III: Concentration: Biology, Chemistry, Earth and Planetary Science, or Physics	50	3 (focused)
Total Items (for each concentration)	**166**	**7**

Note: A candidate must complete all parts of Subtests I and II, but need only complete Subtest III in his/her area of concentration.

Table 2: Subtest Structure of the New CSET: Science (Specialized)

CSET: Science (Specialized) Subtests	Number of Multiple-Choice Items per Test Form	Number of Constructed-Response Items per Test Form
III: Concentration: Biology, Chemistry, Earth and Planetary Science, or Physics	50	3 (focused)
IV: General Science: Biology, Chemistry, Earth and Planetary Science, or Physics	40	1 (focused)
Total Items (for each concentration)	**90**	**4** (focused)

Note: A candidate must complete both Subtests III and IV in his/her area of concentration.

Part II: Standard Setting Studies

Independent panels for each subject area (see Appendix C) were convened to participate in Standard Setting Studies for the new examinations of the CSET Science (Specialized). These Standard Setting Studies were held in Sacramento, CA on October 15-16, 2003, and November 12-13, 2003. The purpose of the standard setting procedure is to provide the Commission recommendations relevant to the determination of passing standards for the examinations, based on the informed judgments of California educators. A total of 35 panel members participated in the standard setting studies. The panels members, who were each appointed by the Executive Director of the Commission after a review of a submitted application, were California educators from across the state, and included curriculum specialists, public school teachers, teacher educators, school administrators, and mentor teachers. Many of the panel members who participated in the CSET: Science (Specialized) standard setting activities had also been a part of the standard setting panels for the CSET: Science examinations in March 2003.

In its April 2003 meeting, the Commission adopted passing standards for the CSET: Multiple Subjects, English, Mathematics, Science, and Social Science. It was determined at that time that candidates must attain minimum passing scores for each subtest of the CSET in a subject area in order to meet the subject matter requirement for the credential. This minimum passing criteria will be used for the CSET: Science (Specialized). In this case, the CSET Science (Specialized) consists of two subtests per specialized area:

Subtest III: which assesses the specific science content knowledge (i.e., biology, chemistry, Earth and planetary science, and physics), and

Subtest IV: which assesses additional general science content knowledge that is specific to each specialized area.

Since passing standards for the Science Subtests III were already established, the panels convened in October an November needed only to focus on recommending passing standards for the new CSET Science (Specialized) Subtest IV.

The standard setting studies began with an orientation and training session, during which panel members received updated information on the new credentials and purpose of this standard setting activity. During the training, they were asked to consider the "just acceptable" candidate. Although many of the examinees will exceed the level of knowledge and skills of the acceptably qualified candidate, none should fall below that level. For this reason, panel members were trained to make judgments based on candidates just at the level of knowledge and skills required of an entry-level teacher candidate to successfully satisfy the subject matter requirement.

To help the panel members become familiar with the examinations, the knowledge and skills associated with the items, and the perspective of the examinee, panel members were provided with a copy of the Subtest Description and the Subtest IV Examination for their field that was administered on September 20, 2003. Under test-like conditions, panel members were asked to read and answer each item independently, and then to score their own performance on the multiple-choice items.

After the simulated test taking and extensive training, panel members were asked to complete three rounds of standard setting tasks based on the test structures. The three rounds of standard setting procedures are described in further detail below.

Round One Standard Setting Ratings

In Round One, panel members independently provided item-by-item ratings, first for the multiple-choice items and then for the constructed-response items.

Multiple-Choice Items

For Round One, panel members were provided the following materials:

- the subtest description;
- the subtest form used for the September 2003 test administration;
- the accompanying subtest form answer keys;
- the Round One Rating Form for multiple-choice items; and
- if appropriate, the item statistics displaying the percent of examinees who answered each test item correctly.

Round One began with a set of approximately ten practice, multiple-choice items for each panel member to rate. This set of items represented a range of item difficulties. Panel members were asked to rate each item by responding to the question that follows.

> *Imagine a hypothetical group of candidates for the Single Subject Teaching Credential in (SCIENCE SPECIALTY), each of whom is just at the level of knowledge and skills important for effective job performance as a beginning teacher in a departmentalized classroom in California public schools.*
>
> *What percent of this group would answer the item correctly**?***

0% - 10% = 1	51% - 60% = 6
11% - 20% = 2	61% - 70% = 7
21% - 30% = 3	71% - 80% = 8
31% - 40% = 4	81% - 90% = 9
41% - 50% = 5	91% - 100% = 10

Panel members were polled as to how they rated each item, and as a panel discussed, when necessary, expected performance of the "just acceptable" candidate and the standard setting procedure. The group also reviewed item statistics (p-values) on each practice test item, where applicable, which provided an indicator of the difficulty level of the item.

Following the practice set, panel members began the same rating process with the multiple-choice items used on the September 20, 2003 operational test forms. NES analyzed the individual and

group results from these item judgments (percentage of "just acceptable" candidates who would answer the item correctly) for use in Round Two of the standard setting process.

Constructed-Response Items

For Round One of the constructed-response item ratings, panel members were provided the following materials:

- the subtest description;
- the subtest form used for the September 2003 test administration;
- the appropriate set of performance characteristics and scoring scale;
- the Subject Matter Advisory Panel-approved marker responses[1] for each score point on the scoring scale; and
- the Round One Rating Form for constructed-response items.

To begin the Round One constructed-response ratings, panel members rated a practice set of two sample items. They were asked to rate each item by responding to the following question.

> *Imagine a hypothetical candidate who is just at the level of knowledge and skills important for effective job performance as a beginning teacher of (SCIENCE SPECIALTY AREA) in California public schools.*
>
> *For this constructed-response item, which of the points on the scoring scale represents the level of response that would be achieved by this individual?*

After the panel had completed the practice set, members were polled for their item ratings, and as a panel discussed the expected level of response in respect to of the concept of the "just acceptable" candidate, the use of the marker responses and scoring scales, and the standard setting procedure.

Following the practice set, panel members began the same rating process with the constructed-response items used on the September 20, 2003 operational test forms. NES analyzed the individual results from these item judgments for use in Round Two of the standard setting process.

Round Two Standard Setting Ratings

Round Two of the standard setting process moved the panels from providing ratings at the item level to ratings made at the component level (i.e., the multiple-choice component and the constructed-response component) for the CSET: Science (Specialized) Subtest IV. Additionally, panel members were asked to provide the percent of points to be allocated for each component in the subtest.

[1] Each Subject Matter Advisory Panel selected responses as marker responses. Marker responses are score-point exemplars used in the training and calibration of scorers.

For Round Two, panel members were provided the following materials:

- the subtest description;
- the Round One Multiple-Choice Item Rating Summary Sheet, which included:
  - the sum of the median rating for each item across all panel members (i.e., the panel's "Computed Median") and;
  - the sum of each panel member's Round One ratings listed in descending order by score value.
- the Round One Constructed-Response Item Rating Summary Sheet, which included:
  - the sum of the median rating for the item across all panel members and doubled to reflect the combined scores examinees receive from two scorers (i.e., the panel's "Computed Median") and;
  - each panel member's Round One constructed-response item rating listed in descending order by score value doubled to reflect the combined scores examinees receive from two scorers.
- the Round Two Subtest component Standard Setting Recommendation Form for multiple choice items; and,
- the Round Two Subtest component Standard Setting Recommendation Form for constructed response items.

(NOTE: Results of individual panel members were provided by identification number only to maintain the confidentiality of each person's ratings.)

Multiple-Choice Items

At this stage in the study, panel members were provided the Round One Item Rating Summary Sheet for discussion and to use that item-level data to inform their component-level recommendations. They were advised that candidates will not "pass" the multiple-choice component alone; a candidate's passing status will be determined at the subtest level, which involves the combination of multiple-choice component and constructed-response component performance.

Panel members worked independently, considering their own aggregated Round One rating and the group median. Each member recommended a single, holistic, Round Two multiple-choice component "cut score" for the subtest, representing the total number of scorable items at the subtest level that would, in his or her judgment, be answered correctly by the "just acceptable" candidate. To make this recommendation, panel members responded to the following question:

Imagine a hypothetical candidate who is just at the level of knowledge and skills important for effective job performance as a beginning teacher of (SCIENCE SPECIALTY AREA) in California public schools.

What is the number of multiple-choice items on the subtest (out of 32 total number of scorable items) that would be answered correctly by this individual?

Constructed-Response Items

For the Round Two constructed-response ratings, panel members were again provided the opportunity to discuss the results of the Round One ratings and the merits of various constructed-response component cut scores at the subtest level. Panel members were reminded that candidates will not "pass" the constructed-response component alone; a candidate's passing status is determined at the subtest level, which is based on the combination of both the multiple-choice and constructed-response components.

Panel members worked independently, considering their own aggregated Round One rating and the group median. Each member recommended a single, holistic, Round Two constructed-response component cut score for the subtest, representing the total number of points at the subtest level that would, in his or her judgment, be obtained by the "just acceptable" candidate. To make this recommendation, panel members responded to the following question:

> *Imagine a hypothetical candidate who is just at the level of knowledge and skills important for effective job performance as a beginning teacher of (SCIENCE SPECIALTY AREA) in California public schools.*
>
> *What is the total score for the constructed-response items on the subtest (out of 6 total number of score points) that would be obtained by this individual?*

Combined Component Scores

Panel members were provided two alternative rules for allocating points consistent with psychometric standards and the structure of each examination: a) multiple-choice component weighting of 80% and constructed-response component weighting of 20%; or b) multiple-choice component weighting of 70% and constructed-response component weighting of 30%. These two options are intended to yield reliable results and are psychometrically defensible. In orienting the panels to this task, they were asked to consider issues of reliability, the length of each component, and the nature of the information provided by each component. Panel members discussed the score combination rules and the rationales for each. Following this discussion, panel members independently made recommendations by responding to the following question:

> *In combining scores on the multiple-choice component and the constructed-response component to yield a total subtest score, what percent of points should be allocated to each component?*
>
> *Check one of the following:*
>
> _____ *80% multiple-choice component and 20% constructed-response component*
>
> _____ *70% multiple-choice component and 30% constructed-response component*

Following this combined component score rating activity, NES collected and analyzed the panel members' recommendations and informed the panelists of the results.

Round Three Standard Setting Ratings

The goal of Round Three of the standard setting process was to produce a passing standard recommendation for each CSET: Science (Specialized) Subtest IV and a set of panel-recommended rules for combining scores from the multiple-choice and constructed-response components.

For Round Three, panel members were provided the following materials:

- the subtest descriptions;
- the Round Two Multiple Choice Results Summary Sheet, which included:
  - the panel's "Computed Median";
  - each panel member's Round Two Multiple Choice rating listed in descending order by score value;
- the Round Two Constructed-Response Results Summary Sheet, which included:
  - the panel's "Computed Median";
  - each panel member's Round Two constructed-response item rating;
  - tabulated panel recommendations on component score combinations;
- the Round Three Subtest Standard Setting Recommendation Form;
- if appropriate, the Summary Statistics Report for subtests, which included;.
  - a set of analyses showing the percent of examinees from the first test administration who would pass the subtest, given possible multiple-choice component and constructed-response component raw score combinations for both component combination rules (i.e., 80%/20% and 70%/30%); and
- if appropriate, the demographic (descriptive) information characterizing the sample of examinees that took the subtest at the September 20, 2003 test administration

  (NOTE: The Summary Statistics Report and demographic information is provided for tests in which 15 or more examinees took the subtest, i.e., biology only)

These materials helped to facilitate a discussion among each panel about their ratings, the nature of the examinee sample, the options for combining component scores, the goal of Round Three, the purpose of the CSET program, and the concept of the just-acceptable candidate.

Panels were cautioned about making judgments based on small numbers of examinees, and were advised that the examinees at the first test administration may or may not reflect the same proportions of all the types and capabilities of examinees in the population that will take the test in the future.

After much discussion, panel members were asked to independently recommend a passing standard and score combination rule for each subtest in their field by responding to the following questions:

Imagine a hypothetical candidate who is just at the level of knowledge and skills important for effective job performance as a beginning teacher of (SCIENCE SPECIALTY AREA) in California public schools.

What is the number of multiple-choice items on the subtest (out of 32 total number of scorable items) that would be answered correctly by this individual?

What is the total score for the constructed-response items on the subtest (out of 6 total number of score points) that would be obtained by this individual?

In combining scores on the multiple-choice component and the constructed-response component to yield a total subtest score, what percent of points should be allocated to each component?

80% multiple-choice component and 20% constructed-response component
OR
70% multiple-choice component and 30% constructed-response component

As the final step to the standard setting studies, each panel member was asked to complete independently a meeting evaluation form regarding the training provided and the task in general.

Results

Following the standard setting studies, NES calculated the median and the distribution of individual Round Three panel recommendations for the multiple-choice and constructed-response test components. Panel recommendations on component score combination rules were also tabulated.

A summary of the panel-based passing score recommendations, including the number of scorable items and the weighting of each component in the total subtest score, is provided in Table 3 on the next page.

Table 3: Panel-Recommended Passing Standards for CSET: Science Subtest IV

CSET: Science Subtest IV	Item Type[1/]	Scorable Items	Possible Score Points	Computed Median based on Panel Recommendations	Component Score Combination Rule [2/]	
					80/20	70/30
• Biology/Life Science	MC	32	32	23.1	✓	
	CR	1	6	4.0		
• Chemistry	MC	32	32	26.0	✓	
	CR	1	6	4.0		
• Earth and Planetary Science	MC	32	32	21.8		✓
	CR	1	6	4.0		
• Physics	MC	32	32	24.7	✓	
	CR	1	6	4.0		

[1/] MC = multiple-choice, CR = constructed-response

[2/] The component score combination rule is formatted as multiple-choice percent/contructed-response percent (e.g., 80/20 is 80% multiple choice / 20% constructed response).

Part III: Staff-Recommended Passing Standards

As described in the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), the standard setting process is a key piece of validity evidence supporting a testing program.

> *Defining the minimum level of knowledge and skill required for licensure or certification is one of the most important and difficult tasks facing those responsible for credentialing. Verifying the appropriateness of the cut score or scores on the tests is a critical element in validity. The validity of the inference drawn from the test depends on whether the standard for passing makes a valid distinction between adequate and inadequate performance. Often, panels of experts are used to specify the level of performance that should be required. Standards must be high enough to protect the public, as well as the practitioner, but not so high as to be unreasonably limiting. Verifying the appropriateness of*

the cut score or scores on a test used for licensure or certification is a critical element of the validity of test results (p.157).

The recommended passing standards the Science (Specialized) Subtests IV are based upon the professional judgments provided by the members of the Science Subject Matter Advisory Panels. Since these panel recommendations are criterion-referenced—based on expert judgment of the minimum required subject matter knowledge for beginning teachers—examinee performance data provides supplemental, though not necessary, information. Due to the limited number of examinees who participated in the initial administrations of these subtests, performance data was not available (see Appendix B for the number of examinees in the initial administration of the CSET: Science (Specialized) Subtests IV). Under these circumstances it is neither practical nor necessary to wait for a minimum number of examinees to set a passing standard.

In making recommendations to the Commission on passing standards for the CSET: Science (Specialized) in biology, chemistry, Earth and planetary science, and physics, staff considered several factors and options that affect the standard setting process. Each consideration is discussed below, followed by staff-recommended passing standards.

Review of Passing Standards

Determining the passing standard(s) for an examination is a careful, conscientious process. For the CSET, it is appropriate to review passing standards periodically to verify that the standards are fulfilling the responsibility of the Commission to award teaching credentials only to those candidates who have fulfilled the subject matter requirement. Since the first administration of these new subtests did not yield 150 examinees, a subsequent passing standard activity will be held to review the passing standards in light of the increased number of examinees. Recommendations for any change in the standards will be presented to the Commission for consideration and adoption.

Standard Error of Measurement

Standard error of measurement is one way to express test reliability and addresses the imprecision of test data. Measurements are not perfectly reliable. In testing, for example, only one score from a single test administration is available for each examinee. An individual examinee's score may, or may not, be accurate. However, the standard error allows us to determine a range within which the examinee's score is likely to lie. Within reasonable limits, the standard error of measurement provides a safeguard against placing undue emphasis on a single numerical score. This is just one index of reliability, and should be applied to the standard setting process in combination with other test-specific characteristics.

Staff-Recommended Passing Standards

Based on these considerations, staff recommends that the Commission adopt the passing standards for the subtests of the CSET forms administered on September 20, 2003 that:

- are equivalent to the raw score points on the multiple-choice component and on the constructed-response component as shown in Table 4;
- are based on the component score combination rules as shown in Table 4; and
- reflect passing standards that are as equivalent as possible for future forms of the test.

The staff-recommended raw score points for multiple-choice and constructed-response components reflect adjustments made for standard errors of measurement as was done in the establishment of the passing standards for the other CSET: Science Subtests I – III in April 2003. Because of the limited number of examinees for these subtests, it is prudent to use the adjustment until such time that there are more examinees so that performance data may be considered. Commission staff will work with NES to schedule subsequent standard setting studies for these subtests, once there are sufficient numbers of examinees to provide the supplemental performance data. The passing standards resulting from these future studies will be brought to the Commission for consideration.

Passing status will be determined on the basis of total subtest performance. Test results will be reported as scaled scores. A scaled score is based on the number of raw score points earned on each component (i.e., multiple-choice and/or constructed-response) and the weighting of each component.

For the CSET, raw scores are converted to a scale from 100 to 300, with a score of 220 representing the passing score as set by the Commission. Scaled scores are used to help ensure that the level of competence required to pass a given test is independent of the particular form of the test taken.

If the Commission adopts the staff-recommended passing standards as indicated in Table 4 on the next page at its December 2003 meeting, NES is currently scheduled to release score reports for the September 2003 test administration and for the November 2003 test administration by December 31, 2003. The next test administration of the CSET is scheduled for January 24, 2004.

Table 4: Staff-Recommended Passing Standards for CSET: Science (Specialized)

CSET: Science (Specialized) Subtest IV – General Science			
	Multiple-Choice Raw Score Points	**Constructed-Response Raw Score Points**	**Component Score Combination Rule MC/CR**
Biology/Life Science	21	4	80/20
Chemistry	24	4	80/20
Earth & Planetary Science	19	4	70/30
Physics	22	4	80/20

With the adoption of passing standards for these subtests, the Commission will have established an examination route for verification of the subject matter requirement for the Single Subject Science (Specialized) credentials in the four science areas of biology, chemistry, Earth and planetary science, and physics.

Appendix A

CSET: SINGLE SUBJECT
THREE-POINT PERFORMANCE CHARACTERISTICS AND SCORING SCALE

Performance Characteristics

PURPOSE	The extent to which the response addresses the constructed-response assignment's charge in relation to relevant CSET subject matter requirements.
SUBJECT MATTER KNOWLEDGE	The application of accurate subject matter knowledge as described in the relevant CSET subject matter requirements.
SUPPORT	The appropriateness and quality of the supporting evidence in relation to relevant CSET subject matter requirements.

SCORING SCALE

SCORE POINT	SCORE POINT DESCRIPTION
3	**The "3" response reflects a command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements.** • The purpose of the assignment is fully achieved. • There is an accurate application of relevant subject matter knowledge. • There is appropriate and specific relevant supporting evidence.
2	**The "2" response reflects a general command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements.** • The purpose of the assignment is largely achieved. • There is a largely accurate application of relevant subject matter knowledge. • There is acceptable relevant supporting evidence.
1	**The "1" response reflects a limited or no command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements.** • The purpose of the assignment is only partially or not achieved. • There is limited or no application of relevant subject matter knowledge. • There is little or no relevant supporting evidence.
U	**The "U" (Unscorable) is assigned to a response that is unrelated to the assignment, illegible, primarily in a language other than English, or does not contain a sufficient amount of original work to score.**
B	**The "B" (Blank) is assigned to a response that is blank.**

Appendix B

CSET: Science (Specialized)
September 20, 2003 Test Administration
Numbers of Examinees by Subtest

CSET: Science (Specialized)	Number of Examinees	Examinees Taking Subtests III & IV
Biology		
Subtest III: Content Specific	237 [1/]	11
Subtest IV: General Science (Biology)	19	
Chemistry		
Subtest III: Content Specific	69 [1/]	7
Subtest IV: General Science (Chemistry)	8	
Earth and Planetary Science		
Subtest III: Content Specific	36 [1/]	1
Subtest IV: General Science (Earth and Planetary Science)	1	
Physics		
Subtest III: Content Specific	41 [1/]	3
Subtest IV: General Science (Physics)	3	

[1/] Subtest III is also used for the regular Single Subject Teaching Credentials in Science.

Appendix C

CSET STANDARD SETTING SCIENCE PANELS

	Biology	Chemistry	Earth and Planetary	Physics	Total
Total Number of Panel Members					
Appointed	20	16	16	16	68
Participated	10	10	6	9	35
Ethnicity					
African American	0	0	0	0	0
Asian	1	0	0	0	1
Hispanic	0	0	0	1	1
White	9	7	6	6	28
Other/Not Provided	0	3	0	2	5
Sex					
Female	3	7	4	3	17
Male	7	3	2	6	18
Region					
North	1	3	2	4	10
South	9	7	4	5	25
Profession					
Public School Educators	8	7	3	7	25
College/University Educators	2	3	3	2	10
Years of Experience					
0-6	1	2	1	0	4
7-10	1	3	0	0	4
11+	8	5	5	9	27
Not Provided	0	0	0	0	0